On The Influences of Variables on Boolean Functions in Product Spaces

Nathan Keller

Einstein Institute of Mathematics, Hebrew University

Jerusalem 91904, Israel

nkeller@math.huji.ac.il

May 26, 2009

Abstract

In this paper we consider the influences of variables on Boolean functions in general product spaces. Unlike the case of functions on the discrete cube where there is a clear definition of influence, in the general case at least three definitions were presented in different papers. We propose a family of definitions for the influence, that contains all the known definitions, as well as other natural definitions, as special cases. We prove a generalization of the BKKKL theorem, which is tight in terms of the definition of influence used in the assertion, and use it to generalize several known results on influences in general product spaces.

1 Introduction

Influences of variables on Boolean functions have been extensively studied during the last few decades. This study led to important applications in Theoretical Computer Science, Combinatorics, Mathematical Physics, Social Choice Theory, and other areas. The basic results on influences were obtained for functions on the discrete cube, but some applications required a generalization of the results to more general product spaces. Unlike the discrete case, where there exists a single natural definition of influence, for general product spaces at least three definitions were presented in different papers. All the definitions are based on dividing the space into subspaces called *fibers*.

Notation 1 *In the following definitions, X denotes a product space $X = X_1 \times X_2 \times \ldots \times X_n$, endowed with a product measure $\mu = \mu_1 \otimes \mu_2 \otimes \ldots \otimes \mu_n$. Throughout the paper, $\log n$ denotes $\log_2 n$.*

Definition 1 *For $x = (x_1, \ldots, x_n) \in X$ and for $1 \leq k \leq n$, the fiber of x in the k-th direction is*

$$s_k(x) = \{y : y_i = x_i, \forall i \neq k\}.$$

The influence of the k-th variable on a function $f : X \to \{0, 1\}$ is the expectation, over all the fibers in the k-th direction, of the influence of the variable on each fiber. The original definition of influences in product spaces, introduced in [5], is the following:

Definition 2 *For $f : X \to \{0, 1\}$, and for $1 \leq k \leq n$, the influence of the k-th variable on f is*

$$I_f(k) = \mu\Big[\{x \in X : f \text{ is non-constant on } s_k(x)\}\Big].$$

Another widely used definition (see, e.g., [11, 13, 16]), is:

Definition 3 *For a function $f : X \to \{0, 1\}$, an element $x = (x_1, \ldots, x_n) \in X$, and for $1 \leq k \leq n$, denote the restriction of f to the fiber $s_k(x)$ by $f_k^x : X_k \to \{0, 1\}$. That is, for all $t \in X_k$,*

$$f_k^x(t) = f(x_1, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_n).$$

The influence of the k-th variable on f is

$$\tilde{I}_f(k) = \mathbb{E}_{x \in X}\mathrm{Var}(f_k^x).$$

The difference between the definitions is demonstrated by the following example:

Example 1 *Consider $X = [0, 1]^n$, endowed with the Lebesgue measure λ. Let $f : [0, 1]^n \to \{0, 1\}$ be defined by:*

$$f(x) = 1 \Longleftrightarrow x_i > 1/n, \forall 1 \leq i \leq n.$$

It is easy to see that for all $1 \leq k \leq n$ we have $I_k(f) \approx 1/e$, while $\tilde{I}_k(f) \approx (n-1)/en^2$.

In general, for every function f we have $\tilde{I}_k(f) \leq I_k(f)/4$, but in many cases $\tilde{I}_k(f)$ is much smaller than $I_k(f)$.

The most well-known theorem concerning influences in general product spaces is the BKKKL theorem [5]:

Theorem 1 (Bourgain, Kahn, Kalai, Katznelson, and Linial) *For every $f : X \to \{0, 1\}$ such that $\mathbb{E}f = p$, there exists a variable k such that $I_f(k) \geq cp(1-p)\log n/n$, where c is a universal constant.*

The BKKKL theorem was used to obtain several important results. For example, it was a central tool in showing that any monotone graph property has a sharp threshold [9].

In this paper we present a generalized definition of influences in general product spaces, which contains Definitions 2 and 3 as special cases:

Definition 4 *Let $h : [0, 1] \to [0, 1]$. For every $f : X \to \{0, 1\}$ and for each $1 \leq k \leq n$, the **h-influence** of the k-th coordinate on f is*

$$I_f^h(k) = \mathbb{E}_{x \in X}h(\mathbb{E}f_k^x).$$

Definition 2 is obtained by substituting $h(t) = 1$ if $t \neq 0, 1$, and $h(t) = 0$ otherwise.[1] Definition 3 is obtained by substituting $h(t) = t(1 - t)$.[2]

---

[1] Here we assume that a function is constant on a fiber if it is constant except for a set of measure zero.

[2] We note that influences toward zero and one (see, e.g., [8]), can be also expressed as h-influences. Influence toward one is obtained by $h(t) = 1 - t$ for $t \neq 0$, and $h(0) = 0$. Influence toward zero is obtained by $h(t) = t$ for $t \neq 1$, and $h(1) = 0$.

We show that the BKKKL theorem can be generalized by replacing the influence $I_k(f)$ with a smaller h-influence, and obtain a full characterization of the h-influences for which the BKKKL theorem holds. As in [5], we prove our results in the case $X = [0,1]^n$, endowed with the Lebesgue measure λ. This case is considered fairly general (see [5, 11]), since many cases of interest can be easily reduced to it.

Theorem 2 *Denote the Entropy function $H(t) = -t \log t - (1-t) \log(1-t)$ by $Ent(t)$. Let $h : [0,1] \to [0,1]$ be a concave function, such that $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$. Then for every $f : [0,1]^n \to \{0,1\}$ with $\mathbb{E}f = p$, there exists $1 \leq k \leq n$ such that the h-influence of the k-th variable on f satisfies*

$$I_f^h(k) \geq cp(1-p) \log n / n,$$

where c is a universal constant.

The advantage of Theorem 2 over the BKKKL theorem is demonstrated by the function $f(x) = 1 \iff (x_i > 1/n, \forall 1 \leq i \leq n)$, presented above. For this function we have $\mathbb{E}(f) \approx 1/e$, and for all k, $I_k(f) \approx 1/e$, and hence the BKKKL theorem is far from being tight in this case. On the other hand, for $h(t) = Ent(t)$ we get $I_k(f) \approx Ent(1/n)/e \approx c' \log n / n$, and thus Theorem 2 is tight in this case.

Furthermore, by examining variants of the *tribes* function presented in [3], we show that Theorem 2 is tight in the following sense:

Proposition 1 *Let $h : [0,1] \to [0,1]$ and let $\epsilon > 0$. If there exists $0 < q < 1$ such that $h(q) \leq \epsilon Ent(q)$, then there exists a function $f : [0,1]^n \to \{0,1\}$ such that $\mathbb{E}f = \Theta(1)$, and for all $1 \leq k \leq n$, the h-influence of the k-th variable on f satisfies $I_f^h(k) \leq c\epsilon \log n / n$, where c is a universal constant.*

The proof of Theorem 2 for monotone functions is simple, and essentially follows the proof of the BKKKL theorem. The reduction from general functions to monotone functions is much more complicated in our case, and involves discretization and convexity arguments. We note that in the special case of functions on the discrete cube endowed with a product measure $\mu = \mu_q^{\otimes n}$, statements similar to Theorem 2 were proved by Friedgut and Kalai ([9], Theorem 3.1), and independently by Talagrand [19].

Using the techniques we develop in the proof of Theorem 2, we provide generalizations of several known results on influences in product spaces. In particular, we generalize a lower bound on the vector of influences obtained by Talagrand [19] for functions on the discrete cube endowed with the product measure $\mu_q^{\otimes n}$. We obtain the following:

Proposition 2 *Let $h : [0,1] \to [0,1]$ be a concave function satisfying $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$. There exists a constant $K > 0$ such that for any function $f : [0,1]^n \to \{0,1\}$ with $\mathbb{E}f = p$,*

$$p(1-p) \leq K \sum_{i \leq n} \frac{I_f^h(i)}{\log \frac{4}{3 I_f^h(i)}}.$$

Our proof uses the proof of Talagrand's result for the case $q = 1/2$, along with the monotonization and discretization technique for h-influences presented in our paper. Since Talagrand's result (for a general q) follows easily from our generalization, our technique can replace the major part of Talagrand's proof (containing a proof of a biased version of Beckner's hypercontractive inequality [2]).

3

This paper is organized as follows: In Section 2 we present the monotonization and discretization techniques for h-influences, which are the main tools used in the paper. The proof of Theorem 2 and its applications are presented in Section 3. In Section 4 we discuss the tightness of our results.

2 Monotonization and Discretization for h-Influences

In this section we generalize to h-influences two central steps of the proof of the BKKKL theorem. The first step is *monotonization*, a shifting technique allowing to replace a Boolean function on the continuous cube by a monotone function with the same expectation and non-higher influences. While for Definition 2 of the influences, the shifting argument is standard and easy, for h-influences the proof is much more complicated. We present this proof, involving discretization and a delicate argument exploiting the concavity of the function h, in Section 2.1. The second step is *discretization*, allowing to replace a monotone function $f : [0,1]^n \to \{0,1\}$ by a function $g : \{0,1\}^m \to \{0,1\}$, where $m = \Theta(n \log n)$, having the following property: The m coordinates can be divided into n equal sets of size $\Theta(\log n)$ each, such that the sum of influences of the coordinates in each set on g is not much higher than the influence of the corresponding coordinate on f. We show that the same can be performed with h-influences, given that for all t, we have $h(t) \geq Ent(t)$. The proof of this claim, presented in Section 2.2, is quite straightforward, and is essentially the same as the proof of Claim 2.8 in [8].

2.1 Monotonization for h-Influences

Definition 5 *A function $f : [0,1]^n \to \mathbb{R}$ is monotone if for all $x, y \in [0,1]^n$,*

$$\forall i (x_i \geq y_i) \Longrightarrow f(x) \geq f(y).$$

An important component in the proofs of the KKL theorem [12], the BKKKL theorem [5] and many other results is *monotonization*. Using the monotonization technique, which is actually a standard shifting argument (see [6]), the examined function is replaced by a monotone function with the same expectation and non-higher influences. The rest of the proof is performed for the monotone function, and the resulting lower bound on its influences yields a lower bound on the influences of the original function.
In this section we prove that the monotonization technique is valid also for h-influences, if h is a concave function. The condition imposed on h holds for all the "natural" definitions of influences in the continuous case, except for Definition 2. However, the validity of the monotonization for this definition was already proved in ([5], Lemma 1).

Theorem 3 *Let $h : [0,1] \to [0,1]$ be a concave function. For every Borel measurable function $f : [0,1]^n \to \{0,1\}$, there exists a monotone function $g : [0,1]^n \to \{0,1\}$, such that:*

  1. $\mathbb{E}f = \mathbb{E}g$.

  2. For all $1 \leq k \leq n$, we have $I_f^h(k) \geq I_g^h(k)$.

Proof The proof of Theorem 3 is composed of three steps.

2.1.1 Reduction to Functions on the Unit Square

The procedure of the monotonization, that allows to arrive to g given f, is the same standard shifting procedure as presented in [5]. The monotonization is performed coordinate-wise. For the k-th coordinate, we consider the fibers in the k-th direction, and for each fiber $s_k(x)$, we replace the function f_k^x by the function

$$Mf_k^x(t) = \begin{cases} 1, & t > 1 - \mathbb{E}(f_k^x) \\ 0, & otherwise. \end{cases}$$

Clearly, the resulting function Mf satisfies $\mathbb{E}(Mf) = \mathbb{E}(f)$. Moreover, it is clear that performing this monotonization procedure in all the coordinates sequentially leads to a monotone function. Hence, the only claim we have to prove is that $I_f^h(j) \geq I_{Mf}^h(j)$ for all $1 \leq j \leq n$, where Mf denotes the monotonization of f in the k-th direction.

For $j = k$ we have $I_f^h(j) = I_{Mf}^h(j)$, since the expectation of f_k^x on each fiber is preserved. Hence, we assume that $j \neq k$. It is sufficient to show that for every fixed value of the $n - 2$ remaining coordinates, the contribution of the family of fibers corresponding to this fixed value to $I_f^h(j)$ is not less than the respective contribution to $I_{Mf}^h(j)$. Therefore, without loss of generality we can fix the remaining coordinates and thus assume that $n = 2$, i.e., $f : [0,1]^2 \to \{0,1\}$.

2.1.2 Proof of a Discrete Version of Theorem 3

Let r be an integer. We assume that the value of f is constant on squares of the form $(\frac{l}{r}, \frac{l+1}{r}) \times (\frac{m}{r}, \frac{m+1}{r})$, for all $0 \leq l, m < r$. Hence, f can be represented by an $r \times r$ matrix A, where $A[i,j]$ denotes the value of $f(x,y)$ for all $(x,y) \in (\frac{i-1}{r}, \frac{i}{r}) \times (\frac{j-1}{r}, \frac{j}{r})$. The monotonization (assuming, without loss of generality, that it is performed according to the rows of the matrix) consists of moving all the 1-s in each row to the rightmost positions in the same row, as described in Figure 1.
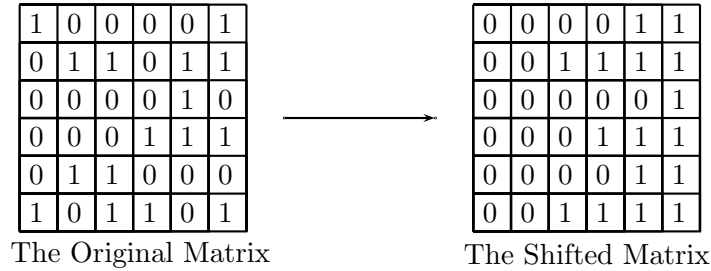
1	0	0	0	0	1
0	1	1	0	1	1
0	0	0	0	1	0
0	0	0	1	1	1
0	1	1	0	0	0
1	0	1	1	0	1

The Original Matrix

$\longrightarrow$

0	0	0	0	1	1
0	0	1	1	1	1
0	0	0	0	0	1
0	0	0	1	1	1
0	0	0	0	1	1
0	0	1	1	1	1

The Shifted Matrix

Figure 1: The Matrices Representing f and Mf

The influence in the "columns" direction is

$$I_f^h(\text{ columns }) = \frac{1}{r} \sum_{j=1}^r h\left(\frac{\sum_{i=1}^r A[i,j]}{r}\right).$$

We show that if h is concave, the influence is not increased by our monotonization. We use two simple observations:

1. If two entire columns are interchanged, the influence does not change.

2. If a '1' is moved from a column with a smaller (or equal) number of 1-s to a column with a bigger (or equal) number, then the influence is not increased.

The first observation is clear. The second observation follows from the concavity of h. Indeed, denote the numbers of 1-s in the columns before the operation by a and b, where $b \geq a$. Since the contribution of all the untouched columns to the influence remains unchanged, we want to prove that

$$h(\frac{b}{r}) + h(\frac{a}{r}) \geq h(\frac{b+1}{r}) + h(\frac{a-1}{r}).$$

There exists $0 < t < 1$ such that $b/r = t(b+1)/r + (1-t)(a-1)/r$. By symmetry, we have $a/r = (1-t)(b+1)/r + t(a-1)/r$. By the concavity of h,

$$h(\frac{b}{r}) = h(t\frac{b+1}{r} + (1-t)\frac{a-1}{r}) \geq th(\frac{b+1}{r}) + (1-t)h(\frac{a-1}{r}),$$

and similarly,

$$h(\frac{a}{r}) \geq (1-t)h(\frac{b+1}{r}) + th(\frac{a-1}{r}).$$

Hence,

$$h(\frac{a}{r}) + h(\frac{b}{r}) \geq h(\frac{b+1}{r}) + h(\frac{a-1}{r}),$$

as asserted.

The transformation from f to Mf can be treated as a concatenation of steps of the two classes discussed in the observations above. Indeed, the transformation can be represented by the following algorithm:

For $i = 0, 1, 2, \ldots, r-1$:

1. Consider columns $1, 2, \ldots, r-i$ of the matrix.

2. Reorder the columns, such that the rightmost column will have the maximal number of 1-s.

3. Consider the zero values in the rightmost column (i.e., column $r-i$). For each such value, look at the corresponding row of the matrix (restricted to the first $r-i$ columns), and if there exist '1' values in the row, interchange one of them with the zero value in the rightmost column.

For example, in Figure 1, the step of the algorithm corresponding to $i = 0$ consists of interchanging the '1' in the cell $A[3,5]$ with the zero in the cell $A[3,6]$, and interchanging the '1' in the cell $A[5,3]$ with the zero in the cell $A[5,6]$. (There is no reordering of the columns since the rightmost column already has the maximal number of 1-s). In the remaining steps of the algorithm, the r-th column is left untouched.

In the output of the algorithm, all the 1-s in each row are concentrated in the rightmost cells. Since during the algorithm, only cells in the same row are interchanged, the output of the algorithm is the matrix representing Mf. The steps of the algorithm are indeed of the two classes discussed in the observations above (reordering the columns and moving a '1' from a column with a smaller (or equal) number of 1-s to a column with a bigger (or equal) number of 1-s). Therefore, the h-influence is not increased by the transformation from f to Mf. This proves the assertion of Theorem 3 for a discrete version of the function f.

2.1.3   Reduction from the General Case to the Discrete Case

We use the following classical result in Measure theory (see, for example, [1], Theorem 5.7):

Theorem 4 *Let μ be a finite measure on a σ-algebra A of subsets of a set Ω, which is generated by an algebra A_0. Then for each $C \in A$, there exists a sequence $\{C_n\}_{n=1}^{\infty}$ such that $C_n \in A_0$ for all n, and*

$$\lim_{n \to \infty} \mu(C_n \triangle C) = 0,$$

where $X \triangle Y = (X \setminus Y) \cup (Y \setminus X)$ is the symmetric difference between the sets X and Y.

Since the algebra generated by the diadic squares (i.e., the squares $(\frac{l}{2^s}, \frac{l+1}{2^s}) \times (\frac{m}{2^s}, \frac{m+1}{2^s})$ for all s and all $0 \le l, m < 2^s$) generates the Borel σ-algebra on the unit square, we get immediately the following corollary:

Corollary 1 *For every Borel measurable function $f : [0,1]^2 \to \{0,1\}$ and every $\epsilon > 0$, there exists an integer s and a function f_s which is constant on squares of the form $(\frac{l}{2^s}, \frac{l+1}{2^s}) \times (\frac{m}{2^s}, \frac{m+1}{2^s})$ for all $0 \le l, m < 2^s$, such that*

$$\int_0^1 \int_0^1 |f - f_s| dx dy \le \epsilon.$$

Let $f : [0,1]^2 \to \{0,1\}$ be a Borel measurable function, and let $\epsilon > 0$. We want to show that $I_f^h(2) \ge I_{Mf}^h(2) - \epsilon$, where Mf denotes the monotonization of f in the first direction (i.e., with respect to the x coordinate). We approximate f by a "discrete" function f_s and use the result on discrete functions obtained in the previous step of the proof.

Since h is concave on $[0,1]$, and thus continuous, there exists δ such that if $|x - y| \le \delta$, then $|h(x) - h(y)| \le \epsilon/4$. By Corollary 1, there exists a function f_s which is constant on squares of the form $(\frac{l}{2^s}, \frac{l+1}{2^s}) \times (\frac{m}{2^s}, \frac{m+1}{2^s})$, such that

$$\int_0^1 \int_0^1 |f(x,y) - f_s(x,y)| dx dy \le \delta\epsilon/4. \tag{1}$$

Let

$$S_1 = \{x \in [0,1] : \int_0^1 |f(x,y) - f_s(x,y)| dy \le \delta\},$$

and $S_2 = [0,1] \setminus S_1$. By the Fubini theorem and Inequality (1), we have $\lambda(S_2) \le \epsilon/4$, where λ is the Lebesgue measure on $\mathbb{R}$. Therefore,

$$\int_{S_2} |h(\int_0^1 f(x,y) dy) - h(\int_0^1 f_s(x,y) dy)| dx \le \epsilon/4. \tag{2}$$

On the other hand, if $x \in S_1$, then

$$|\int_0^1 f(x,y) dy - \int_0^1 f_s(x,y) dy| \le \int_0^1 |f(x,y) - f_s(x,y)| dy \le \delta,$$

and hence,

$$|h(\int_0^1 f(x,y) dy) - h(\int_0^1 f_s(x,y) dy)| \le \epsilon/4.$$

Thus,

$$\int_{S_1} |h(\int_0^1 f(x,y)dy) - h(\int_0^1 f_s(x,y)dy)|dx \le \epsilon/4. \tag{3}$$

Using the triangle inequality and Inequalities (2) and (3), we get:

$$|I_f^h(2) - I_{f_s}^h(2)| = |\int_0^1 h(\int_0^1 f(x,y)dy)dx - \int_0^1 h(\int_0^1 f_s(x,y)dy)dx| \le$$

$$\le \int_0^1 |h(\int_0^1 f(x,y)dy) - h(\int_0^1 f_s(x,y)dy)|dx = \int_{S_1} |h(\int_0^1 f(x,y)dy) - h(\int_0^1 f_s(x,y)dy)|dx +$$

$$+ \int_{S_2} |h(\int_0^1 f(x,y)dy) - h(\int_0^1 f_s(x,y)dy)|dx \le \epsilon/4 + \epsilon/4 = \epsilon/2. \tag{4}$$

In order to bound the term $|I_{Mf}^h(2) - I_{Mf_s}^h(2)|$, we use the following property of the monotonization operation:

Claim 1 *For every pair of measurable functions $f, g : [0,1]^2 \to \{0,1\}$,*

$$\int_0^1 \int_0^1 |Mf(x,y) - Mg(x,y)|dxdy \le \int_0^1 \int_0^1 |f(x,y) - g(x,y)|dxdy,$$

where Mf and Mg represent the monotonizations of f and g with respect to the x coordinate.

Proof   Fix $y \in [0,1]$. Denote $t_0 = 1 - \int_0^1 f(x,y)dx$, and $t_1 = 1 - \int_0^1 g(x,y)dx$. By the definition of Mf and Mg, we have $Mf(x,y) = 0$ if and only if $x < t_0$, and $Mg(x,y) = 0$ if and only if $x < t_1$. Assume w.l.o.g. that $t_0 < t_1$. Then $|Mf(x,y) - Mg(x,y)| = 1$ for $t_0 \le x < t_1$, and $|Mf(x,y) - Mg(x,y)| = 0$ otherwise. Therefore,

$$\int_0^1 |Mf(x,y) - Mg(x,y)|dx = |(1 - \int_0^1 f(x,y)dx) - (1 - \int_0^1 g(x,y)dx)| \le \int_0^1 |f(x,y) - g(x,y)|dx.$$

Hence, by the Fubini theorem,

$$\int_0^1 \int_0^1 |Mf(x,y) - Mg(x,y)|dxdy \le \int_0^1 \int_0^1 |f(x,y) - g(x,y)|dxdy,$$

as asserted. ∎

Combining Claim 1 and Inequality (1) we get:

$$\int_0^1 \int_0^1 |Mf(x,y) - Mf_s(x,y)|dxdy \le \delta\epsilon/4,$$

and hence by the argument applied above to $|I_f^h(2) - I_{f_s}^h(2)|$,

$$|I_{Mf}^h(2) - I_{Mf_s}^h(2)| \le \epsilon/2. \tag{5}$$

Finally, combining Inequalities (4) and (5) with the proof of the theorem in the discrete case, we get:

$$I_f^h(2) - I_{Mf}^h(2) = (I_f^h(2) - I_{f_s}^h(2)) + (I_{f_s}^h(2) - I_{Mf_s}^h(2)) + (I_{Mf_s}^h(2) - I_{Mf}^h(2)) \ge$$

$$\ge (-\epsilon/2) + 0 + (-\epsilon/2) = -\epsilon. \tag{6}$$

This completes the proof of Theorem 3. ∎

2.2 Discretization for h-Influences

Another important component in the proofs of the BKKKL theorem (both the original proof presented in [5] and the simpler proof presented in [8]) is discretization.[3] In this section we follow the discretization suggested in [8], and show that it can be performed not only for Definition 2 of the influences, but also for general h-influences, provided that $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$.

The discretization suggested in [8] is the following. Let $l = 3 \log n$. Subdivide $[0,1]^n$ into 2^{ln} sub-cubes by dividing each of the base intervals into 2^l equal parts. It is shown in [8] that a monotone function $f : [0,1]^n \to \{0,1\}$ can be approximated by a function $\tilde{f}$ that is constant on each of the small sub-cubes. This function naturally corresponds to a function $g : \left(\{0,1\}^l\right)^n \to \{0,1\}$ (by replacing the interval $[m2^{-l}, (m+1)2^{-l}]$ with the binary expansion of m). Each of the initial variables $1 \leq i \leq n$ is now replaced by l variables, $\{i_j\}_{j=1}^l$. It was shown in [5] that for all $1 \leq i \leq n$, we have $\sum_{j=1}^l I_g(i_j) \leq 2I_f(i)$. We prove a similar result for the h-influence, given that h is "big enough".

Proposition 3 *Let $h : [0,1] \to [0,1]$ such that for all $0 \leq t \leq 1$, $h(t) \geq Ent(t)$. Let $f : [0,1]^n \to \{0,1\}$ be a monotone function and let $g : \{0,1\}^{ln} \to \{0,1\}$ be obtained from f by the procedure described above. Then for every $1 \leq i \leq n$,*

$$\sum_{j=1}^l I_g(i_j) \leq 6 I_f^h(i). \tag{7}$$

Proof   We note that it is sufficient to prove the assertion for the contribution of any single fiber $s_i(x)$ to the influences in the two sides of Inequality (7). Let $x \in [0,1]^n$, and consider the function f_i^x, i.e., the restriction of f to the fiber $s_i(x)$. Denote $t_0 = \mathbb{E}(f_i^x)$, and assume that $t_0 \geq 1/2$ (the case $t_0 < 1/2$ is treated later). Since f_i^x is monotone, we have $f_i^x(t) = 0$ if $t < 1 - t_0$ and $f_i^x(t) = 1$ if $t > 1 - t_0$. By the construction of g, the function g_i^x satisfies $g_i^x(m) = 0$ if $m < (1-t_0)2^l$ and $g_i^x(m) = 1$ if $m > (1-t_0)2^l$. Let k satisfy $2^k \leq (1-t_0)2^l < 2^{k+1}$. Then

$$\Pr_{0 \leq m \leq 2^l-1}[g_i^x(m) \neq g_i^x(m \oplus e_j)] \leq \begin{cases} 2^{j+1-l}, & j \leq k \\ 2(1-t_0), & j > k. \end{cases}$$

Indeed, for $j \leq k$, the relation $[g_i^x(m) \neq g_i^x(m \oplus e_j)]$ can be satisfied only for $(1-t_0)2^l - 2^j \leq m \leq (1-t_0)2^l + 2^j$, and hence the probability is bounded from above by 2^{j+1-l}. For $j > k$, the restriction is even stricter: m must satisfy either $0 \leq m \leq (1-t_0)2^l$ or $2^j \leq m \leq (1-t_0)2^l + 2^j$, and hence the probability is bounded by (and actually is equal to) $2(1-t_0)$. Therefore,

$$\sum_j I_{g_i^x}(i_j) \leq \sum_{j \leq k} 2^{j+1-l} + \sum_{j>k} 2(1-t_0) \leq 2^{k+2-l} + 2(1-t_0)(l-1-k) \leq$$

$$\leq 4(1-t_0) + 2(1-t_0)\log\frac{1}{1-t_0} \leq 6(1-t_0)\log\frac{1}{1-t_0} \leq 6Ent(t_0).$$

If $t_0 < 1/2$, then instead of examining the function g_i^x, we consider the dual function $\tilde{g}_i^x(m) = 1 - g_i^x(2^l - 1 - m)$. It is easy to show that $\mathbb{E}(\tilde{g}_i^x) = 1 - t_0$, and that for all j, $I_{\tilde{g}_i^x}(i_j) = I_{g_i^x}(i_j)$. Since $1 - t_0 > 1/2$, we can apply to $\tilde{g}_i^x$ the argument presented above and get

$$\sum_j I_{g_i^x}(i_j) = \sum_j I_{\tilde{g}_i^x}(i_j) \leq 6Ent(1-t_0) = 6Ent(t_0).$$

---

[3] A sketch of another proof of the theorem, that does not use discretization, is presented in [5].

Thus, if $h(t) \geq Ent(t)$ for all t, then for all $x \in [0,1]^n$,

$$\sum_j I_{g_i^x}(i_j) \leq 6 I_{f_i^x}^h(i),$$

as asserted. ∎

Remark 1 *We note that $\overline{I}_f(i) = \sum_j I_g(i_j)$ (where g is as defined above) was treated in [5] as an alternative definition of influences in the continuous case. It follows from the proof of Proposition 3 that this definition is equivalent, up to a multiplicative constant, to h-influence with $h(t) = Ent(t)$. As follows from Theorem 2, this is the "optimal" definition of influence for which the BKKKL theorem holds.*

3   Generalized BKKKL Theorem and Applications

In this section use the monotonization and discretization techniques developed in Section 2 to prove a tight version of the BKKKL theorem for h-influences (Theorem 2), and to generalize several known results concerning influences in product spaces. The results we generalize include:

- A lower bound on the vector of influences obtained by Talagrand [19] (Section 3.2),

- A characterization of functions with a low sum of influences, obtained by Dinur, Friedgut, and Hatami [7, 8, 11] (Section 3.3),

- A relation between the measure of the boundary of a subset of the discrete cube and its influences, obtained by Margulis and Talagrand [15, 18] (Section 3.4), and

- A lower bound on the correlation between monotone subsets of the continuous cube in the average case, obtained in [13] (Section 3.5).

3.1   BKKKL Theorem for h-Influences: Proof of Theorem 2

In the proof of the theorem, we follow the simple proof of the BKKKL theorem presented in [8]. The proof uses the following generalization of the KKL theorem presented in [9]:

Theorem 5 (Friedgut and Kalai) *There exists a constant $c > 0$ such that the following holds: Consider the discrete cube $\{0,1\}^n$ endowed with the uniform measure. Let $f : \{0,1\}^n \rightarrow \{0,1\}$, such that $\mathbb{E}f = p$. If for all k, $I_f(k) \leq \delta$, then*

$$\sum_k I_f(k) \geq cp(1-p)\log(1/\delta).$$

The proof of Theorem 2 is a straightforward combination of Theorem 3, Proposition 3, and Theorem 5, as follows:

Proof of Theorem 2: Let $f : [0,1]^n \rightarrow \{0,1\}$, such that $\mathbb{E}f = p$. We want to find a lower bound on the h-influences of f. By Theorem 3, we can replace f by a monotone function with the same expectation and lower h-influences. A lower bound on the h-influences of the new function implies the same bound on the h-influences on f, and hence we can assume w.l.o.g. that f is monotone. Since $h(t) \geq Ent(t)$ for all t, we can use Proposition 3 to approximate f

by a function $g : \{0,1\}^{nl} \to \{0,1\}$, such that for all i, $\sum_j I_g(i_j) \leq 6I_f^h(i)$. Now assume that all the h-influences of f satisfy $I_f^h(i) \leq c_1 \log n/n$ for some constant c_1. In this case, for all i,j we have $I_g(i_j) \leq 6c_1 \log n/n$, and thus, by Theorem 5,

$$\sum_i I_f^h(i) \geq \frac{1}{6} \sum_{i,j} I_g(i_j) \geq c_2 p(1-p) \log n,$$

for some constant c_2. Therefore, there exists a coordinate i such that $I_f^h(i) \geq c_2 p(1-p) \log n/n$, as asserted. ∎

A slight modification of the proof yields the following generalization of Theorem 5 to h-influences, that shall be used later:

Proposition 4 *Let $h : [0,1] \to [0,1]$ be a concave function satisfying $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$. There exists a constant $c > 0$ such that the following holds: Let $f : [0,1]^n \to \{0,1\}$, such that $\mathbb{E}f = p$. If for all k, $I_f^h(k) \leq \delta$, then*

$$\sum_k I_f^h(k) \geq cp(1-p) \log(1/\delta).$$

Remark 2 *We note that the strongest form of the assertion of Theorem 2 is obtained for the function $h(t) = Ent(t)$. The same holds also for Proposition 4, and for Propositions 5, 6, 8, and 10, presented in the following sections. However, for the sake of generality, we state these propositions for a general h-influence, satisfying the condition "$h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$".*

3.2 Generalization of Talagrand's Lower Bound on the Vector of Influences

In [19], Talagrand proved the following strengthening of the KKL theorem for functions on the discrete cube endowed with a general product measure:

Theorem 6 (Talagrand) *Consider the discrete cube $\{0,1\}^n$ endowed with the product measure μ_q defined by $\mu_q(x) = q^{\sum x_i}(1-q)^{n-\sum x_i}$. There exists a constant $K > 0$ such that for any function $f : \{0,1\}^n \to \{0,1\}$ with $\mathbb{E}f = p$,*

$$p(1-p) \leq Kq(1-q) \log \frac{2}{q(1-q)} \sum_{i \leq n} \frac{I_f(i)}{\log \frac{1}{q(1-q)I_f(i)}}. \tag{8}$$

The proof of Theorem 6 uses the biased Fourier-Walsh expansion for functions on the discrete cube, and a biased version of Beckner's hypercontractive inequality [2], which is proved in [19].[4]

Using the technique presented above, we can generalize Theorem 6 to h-influences: [5]

---

[4]The biased hypercontractive inequality proved in ([19], Lemma 2.1) yields a hypercontractivity constant of $q(1-q)$. As was shown later in [17], the optimal constant is bigger (of order $Ent(q)$). However, this in-optimality affects the assertion of Theorem 6 only in the constant factor K which is not specified in [19].

[5]Another generalization of Theorem 6 to the continuous case was obtained recently by Hatami [11], for a combination of Definitions 2 and 3 of the influences. Hatami's result does not follow from Proposition 5, but also does not imply it.

Proposition 5 *Let $h : [0,1] \to [0,1]$ be a concave function satisfying $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$. There exists a constant $K > 0$ such that for any function $f : [0,1]^n \to \{0,1\}$ with $\mathbb{E}f = p$,*

$$p(1-p) \leq K \sum_{i \leq n} \frac{I_f^h(i)}{\log \frac{4}{3 I_f^h(i)}}. \tag{9}$$

Proof   Let $f : [0,1]^n \to \{0,1\}$. As in the proof of Theorem 2, monotonization and discretization allow to replace f by a function $g : \{0,1\}^{nl} \to \{0,1\}$, such that $\mathbb{E}g = \mathbb{E}f = p$, and for all i, $\sum_j I_g(i_j) \leq 6 I_f^h(i)$. Applying Theorem 6 in the case $q = 1/2$ to g, we get

$$p(1-p) \leq K \sum_{1 \leq i \leq n, 1 \leq j \leq l} \frac{I_g(i_j)}{\log \frac{4}{I_g(i_j)}} \leq 2K \sum_{1 \leq i \leq n, 1 \leq j \leq l} \frac{I_g(i_j)}{\log \frac{8}{I_g(i_j)}}. \tag{10}$$

Note that for all $1 \leq i \leq n$,

$$\sum_{1 \leq j \leq l} \frac{I_g(i_j)}{\log \frac{8}{I_g(i_j)}} \leq \sum_{1 \leq j \leq l} \frac{I_g(i_j)}{\log \frac{8}{\sum_{1 \leq j \leq l} I_g(i_j)}} = \frac{\sum_{1 \leq j \leq l} I_g(i_j)}{\log \frac{8}{\sum_{1 \leq j \leq l} I_g(i_j)}}.$$

Since the function $\varphi(x) = x/\log(8/x)$ is monotone increasing in x in $[0, 8)$, we get

$$\sum_{1 \leq i \leq n} \sum_{1 \leq j \leq l} \frac{I_g(i_j)}{\log \frac{8}{I_g(i_j)}} \leq \sum_{1 \leq i \leq n} \frac{\sum_{1 \leq j \leq l} I_g(i_j)}{\log \frac{8}{\sum_{1 \leq j \leq l} I_g(i_j)}} \leq \sum_{1 \leq i \leq n} \frac{6 I_f^h(i)}{\log \frac{8}{6 I_f^h(i)}}.$$

Finally, substitution into Inequality (10) yields:

$$p(1-p) \leq 2K \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq l} \frac{I_g(i_j)}{\log \frac{8}{I_g(i_j)}} \leq 2K \sum_{1 \leq i \leq n} \frac{6 I_f^h(i)}{\log \frac{8}{6 I_f^h(i)}} = K' \sum_{1 \leq i \leq n} \frac{I_f^h(i)}{\log \frac{4}{3 I_f^h(i)}},$$

as asserted. ∎

We note that Theorem 6 (for a general q) follows from Proposition 5, using a standard transformation from the biased measure on the discrete cube to the Lebesgue measure on the continuous cube. Indeed, consider the discrete cube $\{0,1\}^n$ endowed with the measure μ_q. Define $G : [0,1] \to \{0,1\}$ by $G(x) = 0$ if $x \leq 1 - q$, and $G(x) = 1$ if $x > 1 - q$. For a function $f : \{0,1\}^n \to \{0,1\}$, define $\tilde{f} : [0,1]^n \to \{0,1\}$ by

$$\tilde{f}(x_1, x_2, \ldots, x_n) = f(G(x_1), G(x_2), \ldots, G(x_n)).$$

It is easy to see that $\mathbb{E}\tilde{f} = \mathbb{E}f$, where the expectation in the left hand side is w.r.t. the Lebesgue measure on the continuous cube, and the expectation in the right hand side is w.r.t. the measure μ_q on the discrete cube. Applying Proposition 5 to the function $\tilde{f}$, we get

$$p(1-p) \leq K \sum_{1 \leq i \leq n} \frac{I_{\tilde{f}}^h(i)}{\log \frac{4}{3 I_{\tilde{f}}^h(i)}}, \tag{11}$$

where $p = \mathbb{E}\tilde{f} = \mathbb{E}f$. Due to the construction of $\tilde{f}$, the contribution of each non-constant fiber to an influence of $\tilde{f}$ is $h(q)$, and hence taking $h(t) = Ent(t)$, the contribution of each non-constant fiber is $Ent(q)$. Thus, for all $1 \leq i \leq n$, we have:

$$I_{\tilde{f}}^h(i) = Ent(q) I_f(i),$$

12

where the influence in the right hand side is w.r.t. the measure μ_q on the discrete cube. Substituting into Inequality (11), we get:

$$p(1-p) \le KEnt(q) \sum_{1 \le i \le n} \frac{I_f(i)}{\log \frac{4}{3Ent(q)I_f(i)}}. \tag{12}$$

Finally, there exist constants K_1 and K_2 such that for all q and all $I_f(i)$,

$$Ent(q) \le K_1 q(1-q) \log \frac{2}{q(1-q)},$$

and

$$\log \frac{4}{3Ent(q)I_f(i)} \ge K_2 \log \frac{1}{q(1-q)I_f(i)}.$$

Therefore, the assertion of Theorem 6 follows from Inequality (12).

This shows that the biased hypercontractive inequality used in the proof in [19] can be replaced by the original hypercontractive inequality (i.e. the inequality for the uniform measure), combined with our argument presented above.

3.3 Functions with a Low Sum of Influences

One of the most useful results concerning influences of variables on functions on the discrete cube is the following theorem, due to Friedgut [7], asserting that if the sum of influences is small then the function essentially depends on a few coordinates.

Theorem 7 (Friedgut) *Consider $\{0,1\}^n$ as a measure space with the uniform measure. Let $f : \{0,1\}^n \to \{0,1\}$ such that $\sum_k I_f(k) = t$, and let $\epsilon > 0$. Denote $M = t/\epsilon$. There exists a Boolean function g depending only on $\exp((2+\sqrt{\frac{2\log(4M)}{M}})M)$ variables, such that $\Pr[f \ne g] \le \epsilon$.*

Dinur and Friedgut [8] observed that using discretization, Theorem 7 can be generalized to monotone functions on the continuous cube:

Theorem 8 (Dinur and Friedgut) *There exists a constant $c > 0$ such that the following holds: Let $\epsilon > 0$, and let $f : [0,1]^n \to \{0,1\}$ be a monotone function. If $\sum_i I_f(i) \le B$, then there exists a set $J \subset \{1,2,\ldots,n\}$ with $|J| \le \exp(cB/\epsilon)$ and a function $g : [0,1]^n \to \{0,1\}$ depending only on the coordinates in J, such that $\|f - g\|_2^2 \le \epsilon$.*

Dinur and Friedgut [8] conjectured that the assertion of Theorem 8 holds even without the monotonicity assumption. This conjecture was disproved by Hatami [11]. On the other hand, Hatami proved (for general Boolean functions on the continuous cube) that if the sum of influences of f is small, then f can be approximated by a function having a *decision tree of bounded depth* (see [11] for the definitions).

Theorem 9 (Hatami) *Let $f : [0,1]^n \to \{0,1\}$ satisfy $\sum_i I_f(i) \le B$. Then for every $\epsilon > 0$, there exists a function $g : [0,1]^n \to \{0,1\}$ such that $\|f - g\|_2^2 \le \epsilon$, and g has a decision tree of depth at most $\exp(cB/\epsilon^2)$, where c is a universal constant.*

Using the techniques presented above we can generalize Theorems 8 and 9 to h-influences.

Proposition 6 *Let $h : [0,1] \to [0,1]$ satisfy $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$. There exists a constant $c > 0$ such that the following holds: Let $\epsilon > 0$, and let $f : [0,1]^n \to \{0,1\}$ be a monotone function. If $\sum_i I_f^h(i) \leq B$, then there exists a set $J \subset \{1, 2, \ldots, n\}$ with $|J| \leq \exp(cB/\epsilon)$ and a function $g : [0,1]^n \to \{0,1\}$ depending only on the coordinates in J, such that $||f - g||_2^2 \leq \epsilon$.*

Proposition 6 follows immediately from Theorem 7 using Proposition 3. As in Section 3.2, we can apply Proposition 6 with $h(t) = Ent(t)$ to functions on the discrete cube endowed with the measure μ_q to get:

Proposition 7 *There exists an absolute constant $c > 0$ such that the following holds: Consider the discrete cube $\{0,1\}^n$ endowed with the measure μ_q. If a monotone function $f : \{0,1\}^n \to \{0,1\}$ satisfies $\sum_i I_f(i) \leq B$, then there exists $g : \{0,1\}^n \to \{0,1\}^n$ depending on at most $\exp(cEnt(q)B/\epsilon)$ coordinates, such that $||f - g||_2^2 \leq \epsilon$.*

For monotone functions, Proposition 7 gives a more precise result than Theorem 4.1 in [7] (in which the exact dependence on q was not specified).

The generalization of Theorem 9 is as follows:

Proposition 8 *Let $h : [0,1] \to [0,1]$ be a concave function satisfying $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$. Let $f : [0,1]^n \to \{0,1\}$ satisfy $\sum_i I_f^h(i) \leq B$. Then for every $\epsilon > 0$, there exists a function $g : [0,1]^n \to \{0,1\}$ such that $||f - g||_2^2 \leq \epsilon$, and g has a decision tree of depth at most $\exp(cB/\epsilon^2)$, where c is a universal constant.*

The proof of Proposition 8 is a minor modification of the proof of Theorem 9 presented in [11]. The only two changes are replacing ordinary influences with h-influences throughout the proof, and replacing "Theorem B" used in the proof by the following proposition, which follows immediately from Proposition 4.

Proposition 9 *Let $h : [0,1] \to [0,1]$ be a concave function satisfying $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$. For all $f : [0,1]^n \to \{0,1\}$ with $\mathbb{E}f = p$, there exists $1 \leq i \leq n$, such that*

$$I_f^h(i) \geq e^{-\frac{c}{p(1-p)} \sum_{j=1}^n I_f^h(j)},$$

where c is a universal constant.

In Section 4 we show that for some cases of interest, the advantage of Proposition 6 over Theorem 8 is significant.

3.4 A Relation Between the Sum of Influences and the Size of the Boundary

Definition 6 *Consider a product space $X = X_1 \times X_2 \times \ldots \times X_n$ endowed with a product measure $\mu = \mu_1 \otimes \ldots \otimes \mu_n$, and let $A \subset X$. The boundary of A is*

$$\partial A = \{x \in A : \exists (1 \leq i \leq n), \text{ s.t. the function } 1_A \text{ is non-constant on the fiber } s_i(x)\}.$$

In [15], Margulis proved that for subsets of the discrete cube endowed with the uniform measure, the size of the boundary and the sum of influences cannot be small simultaneously. For monotone subsets of the discrete cube, Talagrand [18] gave the following precise form to this statement:

Theorem 10 (Talagrand) *Consider the discrete cube $\{0,1\}^n$ endowed with the uniform measure μ. For any monotone subset $A \subset \{0,1\}^n$ with $\mu(A) \leq 1/2$,*

$$\mu(\partial A) \sum_i I_{1_A}(i) \geq c\mu(A)^2 \log \frac{e}{\mu(A)},$$

where c is a universal constant.

Theorem 10 can be generalized to subsets of the continuous cube, as follows:

Proposition 10 *Let $h : [0,1] \to [0,1]$ be a function satisfying $h(t) \geq Ent(t)$ for all $0 \leq t \leq 1$. Consider the continuous cube $[0,1]^n$ endowed with the Lebesgue measure λ. For any monotone subset $A \subset [0,1]^n$ with $\lambda(A) \leq 1/2$,*

$$\lambda(\partial A) \sum_i I_{1_A}^h(i) \geq c\lambda(A)^2 \log \frac{e}{\lambda(A)}, \tag{13}$$

where c is a universal constant.

Proof   As in the proof of Theorem 2, discretization allows to replace the function 1_A by a function $g : \{0,1\}^{nl} \to \{0,1\}$, such that $\mathbb{E}g = \mathbb{E}1_A$, and for all i, $\sum_j I_g(i_j) \leq 6I_{1_A}^h(i)$. Denote the subset of the discrete cube corresponding to g by B. Applying Theorem 10 to B, we get

$$\mu(\partial B) \sum_{i,j} I_g(i_j) \geq c\mu(B)^2 \log \frac{e}{\mu(B)}. \tag{14}$$

By the construction, we have $\sum_{i,j} I_g(i_j) \leq 6 \sum_i I_{1_A}^h(i)$, and $\mu(B) = \lambda(A)$. In addition, we observe that $\mu(\partial B) \leq \lambda(\partial A)$. Indeed, if $x \in \partial B$ then there exist i, j such that either ($x \in B$ and $x \oplus e_{i_j} \notin B$) or ($x \notin B$ and $x \oplus e_{i_j} \in B$). Recall that the pre-image of x (in the discretization procedure) is a sub-cube of the continuous cube, denoted by A_x, such that $\lambda(A_x) = 2^{-nl}$. For all $y \in A_x$, the function 1_A is non-constant on the fiber $s_i(y)$, and hence $A_x \subset \partial A$. Hence, $\mu(\partial B) \leq \lambda(\partial A)$. The assertion of Proposition 10 (with the constant $c/6$) follows now immediately from Inequality (14). ∎

We note that a stronger version of Theorem 10 was proved by Talagrand in [21]. It seems challenging to find a generalization of this version to the continuous cube.

3.5   Lower Bound on the Correlation Between Monotone Subsets of the Continuous Cube

One of the most well-known correlation inequalities for monotone subsets of the discrete cube is the following inequality, due to Harris [10] and Kleitman [14]:

Theorem 11 (Harris, Kleitman) *Let A, B be monotone subsets of $\{0,1\}^n$ endowed with the uniform measure μ. Then*

$$\mu(A \cap B) \geq \mu(A)\mu(B), \tag{15}$$

i.e., the correlation of A and B is nonnegative.

In [20], Talagrand established a lower bound on the correlation in terms of how much the two sets depend simultaneously on the same coordinates.

Theorem 12 (Talagrand) *Let A, B be monotone subsets of $\{0, 1\}^n$ endowed with the uniform measure μ. Then*

$$\mu(A \cap B) - \mu(A)\mu(B) \geq K\varphi(\sum_{i \leq n} I_{1_A}(i)I_{1_B}(i)), \tag{16}$$

where $\varphi(x) = x/\log(e/x)$ and K is a universal constant.

In [13] it was shown that when the correlation is averaged amongst all $A, B \in T$ for any family T of monotone subsets of $\{0, 1\}^n$, the lower bound asserted in Theorem 12 can be improved.

Proposition 11 ([13]) *Let T be a family of monotone subsets of the discrete cube endowed with the uniform measure μ. Then*

$$\sum_{A,B \in T} (\mu(A \cap B) - \mu(A)\mu(B)) \geq \frac{1}{4} \sum_{A,B \in T} \sum_{i \leq n} I_{1_A}(i)I_{1_B}(i). \tag{17}$$

Furthermore, it was shown in [13] that Proposition 11 can be generalized to the correlation of monotone subsets of the continuous cube, using Definition 3 of the influences in the continuous case.

Proposition 12 ([13]) *Let T be a family of monotone subsets of the continuous cube endowed with the Lebesgue measure λ. Then*

$$\sum_{A,B \in T} (\lambda(A \cap B) - \lambda(A)\lambda(B)) \geq 3 \sum_{A,B \in T} \sum_{i \leq n} \tilde{I}_{1_A}(i)\tilde{I}_{1_B}(i). \tag{18}$$

It was also shown in [13] that the natural generalization of Proposition 11 does not hold under Definition 2 of the influences in the continuous case.

Using h-influences, the assertion of Proposition 12 can be generalized significantly:

Proposition 13 *Let $1/2 < \alpha \leq 1$, and let $h_\alpha : [0, 1] \to [0, 1]$ be defined by $h_\alpha(t) = [t(1 - t)]^\alpha$. There exists $c = c(\alpha)$ such that for any family T of monotone subsets of the continuous cube endowed with the Lebesgue measure λ,*

$$\sum_{A,B \in T} (\lambda(A \cap B) - \lambda(A)\lambda(B)) \geq c(\alpha) \sum_{A,B \in T} \sum_{i \leq n} I_{1_A}^{h_\alpha}(i)I_{1_B}^{h_\alpha}(i). \tag{19}$$

Proof The proof is similar to the proof of Proposition 12, presented in [13]. The main observation in the proof is that for $1/2 < \alpha \leq 1$, the influences $I_{1_A}^{h_\alpha}(i)$ can be represented as the Fourier coefficients of the function 1_A with respect to an appropriate orthonormal system of functions in $L^2([0, 1])$. Consider the functions $r_i' : (0, 1) \to \mathbb{R}$ defined by:

$$r_i'(t) = \alpha(2t - 1)[t(1 - t)]^{\alpha - 1}.$$

Since $\int_0^1 r_i'(t)dt = 0$, the functions $\{r_i'\}_{i=1}^n$ along with the constant function constitute an orthogonal system in $L^2([0, 1])$. Let

$$c(\alpha) = \left(\int_0^1 (\alpha(2t - 1)[t(1 - t)]^{\alpha - 1})^2 dt \right)^{-1}.$$

Note that the integral is finite since $\alpha > 1/2$. The functions $r_i(t) = c(\alpha)^{1/2} r_i'(t)$ satisfy $\int_0^1 r_i^2(t) dt = 1$, and hence the system $\{r_i\}_{i=1}^n$ along with the constant function $r_\emptyset \equiv 1$ is an orthonormal system of functions in $L^2([0,1])$. By ([13], Lemma 7), this implies that for any family T of monotone subsets of the continuous cube,

$$\sum_{A,B \in T} (\lambda(A \cap B) - \lambda(A)\lambda(B)) \geq \sum_{A,B \in T} \sum_{i \leq n} \int_{[0,1]^n} 1_A r_i(x) d\lambda(x) \int_{[0,1]^n} 1_B r_i(x) d\lambda(x). \qquad (20)$$

We observe that for all $1 \leq i \leq n$,

$$\int_{[0,1]^n} 1_A r_i(x) d\lambda(x) = c(\alpha)^{1/2} I_{1_A}^{h_\alpha},$$

and similarly for 1_B. Indeed, consider a single fiber $s_i(x)$, and denote the restriction of the function 1_A to the fiber $s_i(x)$ by $f_i^x : [0,1] \to \{0,1\}$. Since the function f_i^x is Boolean and monotone, there exists t_0 such that for all $t < t_0$, $f_i^x(t) = 0$, and for all $t > t_0$, $f_i^x(t) = 1$. In this case we have

$$\int_0^1 f_i^x r_i(t) dt = \int_{t_0}^1 c(\alpha)^{1/2} \alpha(2t-1)[t(1-t)]^{\alpha-1} dt = c(\alpha)^{1/2}[t_0(1-t_0)]^\alpha = c(\alpha)^{1/2} h_\alpha(\mathbb{E} f_i^x).$$

Hence,

$$\int_{[0,1]^n} 1_A r_i(x) d\lambda(x) = \mathbb{E}_{x \in [0,1]^n} \int_0^1 f_i^x r_i(t) dt = c(\alpha)^{1/2} \mathbb{E}_{x \in [0,1]^n} h_\alpha(\mathbb{E} f_i^x) = c(\alpha)^{1/2} I_{1_A}^{h_\alpha}(i).$$

Substituting into Inequality (20), the assertion of the proposition follows. ∎

The following example demonstrates that the natural generalization of Proposition 11 does not hold for h-influences of the class $h_\alpha(t) = [t(1-t)]^\alpha$ with $\alpha < 1/2$.

Example 2 *Let $A \subset [0,1]^n$ be defined by*

$$A = \{x \in [0,1]^n : (\forall i : x_i \geq 1/n)\}.$$

Clearly, $\lambda(A) = (1-1/n)^n \approx 1/e$, and hence,

$$\lambda(A \cap A) - \lambda(A)\lambda(A) \approx (1/e) - (1/e)^2 = \Theta(1).$$

The function 1_A is non-constant on a fiber $s_k(x)$ if and only if $x_j \geq 1/n$ for all $j \neq k$. Hence, the h_α-influence of the k-th coordinate on 1_A is

$$I_{1_A}^{h_\alpha}(k) = (1-1/n)^{n-1} \left(\frac{n-1}{n^2}\right)^\alpha = \Theta(n^{-\alpha}),$$

and thus,

$$\sum_k I_{1_A}^{h_\alpha}(k) I_{1_A}^{h_\alpha}(k) = \Theta(n^{1-2\alpha}).$$

Therefore, if $\alpha < 1/2$ then the assertion of Proposition 13 is far from being correct, even for the self-correlation of the set A.

If the assertion of Proposition 13 holds for $\alpha = 1/2$, then the above example shows its tightness. However, we weren't able to prove the assertion in this case. It seems challenging to find a generalization of Talagrand's Theorem 12 to the continuous cube using h-influences, but it seems that the techniques of the current paper cannot provide such result.

4 Tightness of Results

The tightness of most of our results can be shown using a biased variant of the *tribes* function presented in [3]. We note that this function was already used to show the tightness of the results in [7, 9, 19]. In the construction below we assume that $q \leq 1/2$. The case $q > 1/2$ is treated in the end of this section.

4.1 Tightness of Theorem 2

Consider the discrete cube $\{0,1\}^n$ endowed with the product measure μ_q. Partition the set $\{1,2,\ldots,n\}$ into sets $\{T_1, T_2, \ldots, T_{n/r}\}$ of size

$$r = \frac{\log n - \log\log n + \log\log(1/q)}{\log(1/q)}$$

each, and define $f : \{0,1\}^n \to \{0,1\}$ by setting $f(x) = 1$ if and only if there exists i such that $x_j = 1$ for all $j \in T_i$. This function can be transformed to a function on the continuous cube as follows: Define $G : [0,1] \to \{0,1\}$ by $G(x) = 0$ if $x \leq 1 - q$, and $G(x) = 1$ if $x > 1 - q$, and let $\tilde{f} : [0,1]^n \to \{0,1\}$ be defined by

$$\tilde{f}(x_1, x_2, \ldots, x_n) = f(G(x_1), G(x_2), \ldots, G(x_n)).$$

It is easy to see that

$$\mathbb{E}\tilde{f} = 1 - (1 - q^r)^{n/r} \approx 1 - 1/e,$$

where the expectation is with respect to the Lebesgue measure on the continuous cube. The function $\tilde{f}$ is non-constant on a fiber $s_k(x)$ where $k \in T_i$ if for all $j \in T_i$ we have $x_j > 1 - q$, and for each $i' \neq i$, there exists $j \in T_{i'}$ such that $x_j \leq 1 - q$. In each of the non-constant fibers we have $\mathbb{E}\tilde{f}_k^x = q$. Thus, for all $h : [0,1] \to [0,1]$ and for all $1 \leq k \leq n$,

$$I_{\tilde{f}}^h(k) = q^{r-1}(1 - q^r)^{(n/r)-1}h(q) \approx \frac{\log n}{(en)q\log(1/q)}h(q) \leq \frac{2\log n}{(en)Ent(q)}h(q),$$

where the last inequality holds since for all $q \leq 1/2$ we have $q\log(1/q) \geq (1-q)\log(1/(1-q))$. This example shows that the assertion of Theorem 2 is tight:

Proposition 14 *Let $h : [0,1] \to [0,1]$ and let $\epsilon > 0$. If there exists $0 < q < 1$ such that $h(q) \leq \epsilon Ent(q)$, then there exists a function $f : [0,1]^n \to \{0,1\}$ such that $\mathbb{E}f = \Theta(1)$, and for all $1 \leq k \leq n$, the h-influence of the k-th variable on f satisfies $I_f^h(k) \leq c\epsilon \log n/n$, where c is a universal constant.*

4.2 Tightness of Proposition 5

The same example shows the tightness of Proposition 5. Indeed, if there exists q such that $h(q) \leq \epsilon Ent(q)$, then for the corresponding function $\tilde{f}$, we have

$$I_{\tilde{f}}^h(k) \leq \frac{2\epsilon \log n}{en},$$

for all $1 \leq k \leq n$. Since the function $g(x) = x/\log(4/3x)$ is monotone increasing in x, we have

$$\sum_{k=1}^n \frac{I_{\tilde{f}}^h(k)}{\log \frac{4}{3I_{\tilde{f}}^h(k)}} = \sum_{k=1}^n g(I_{\tilde{f}}^h(k)) \leq ng\left(\frac{2\epsilon \log n}{en}\right) \approx c\frac{\epsilon \log n}{\log(n/\epsilon)} \leq c'\epsilon,$$

where the last inequality holds for $n \gg 1/\epsilon$. This contradicts Inequality (9) since for $\tilde{f}$, the left hand side of the inequality is $\Theta(1)$.

4.3 Tightness of Proposition 6

In order to show the advantage of Proposition 6 over Theorem 8, we consider a slight modification of the tribes function examined above. For $0 < q \leq 1/2$ and for $m < n$, we construct the tribes function $\tilde{f} : [0,1]^m \rightarrow \{0,1\}$, as above. Then we define $\overline{f} : [0,1]^n \rightarrow \{0,1\}$ by $\overline{f}(x_1, x_2, \ldots, x_n) = \tilde{f}(x_1, x_2, \ldots, x_m)$. Clearly, $\overline{f}$ depends on m variables. By the calculation presented above, we have

$$\sum_k I_{\overline{f}}^h(k) = \sum_k I_{\tilde{f}}^h(k) \leq \frac{2\log m}{eEnt(q)} h(q),$$

and hence, for $h(t) = Ent(t)$, we get $\sum_k I_{\overline{f}}^h(k) \leq 2\log m/e$. Therefore, Proposition 6 implies that $\overline{f}$ can be approximated by a function depending on at most $\exp(c\log m/\epsilon)$ variables, which is tight, up to the (c/ϵ) factor. For comparison, note that $\sum_k I_{\overline{f}}(k) \approx m/e$, and hence Theorem 8 implies approximation by a function depending on $\exp(cm/\epsilon)$ variables, which is far from being tight.

4.4 Tightness of Proposition 10

Finally, in order to show the tightness of Proposition 10, we consider a balanced threshold function on the biased discrete cube. Consider the discrete cube $\{0,1\}^n$ endowed with the measure μ_q, and let

$$A = \{x \in \{0,1\}^n : \sum_i x_i > \lfloor nq \rfloor\}.$$

Using the function G defined above, A can be transformed to $\tilde{A} \subset [0,1]^n$. It is well-known that $\lambda(\tilde{A}) = \mu_q(A) = \Theta(1)$. We have

$$\partial A = \{x \in \{0,1\}^n : \lfloor nq \rfloor \leq \sum_i x_i \leq \lfloor nq \rfloor + 1\},$$

and hence it can be shown using Stirling's formula that

$$\lambda(\partial\tilde{A}) = \mu_q(\partial A) \approx \frac{2}{\sqrt{2\pi nq(1-q)}}.$$

The function $1_{\tilde{A}}$ can be non-constant on the fiber $s_k(x)$ only if $x \in \partial\tilde{A}$, and the expectation of the function on each non-constant fiber is q. Thus,

$$\sum_{i=1}^n I_{1_{\tilde{A}}}^h(i) \leq n\mu_q(\partial A)h(q) \approx \frac{2n}{\sqrt{2\pi nq(1-q)}} h(q).$$

Taking $h(q) = Ent(q)$, we get

$$\lambda(\partial\tilde{A}) \sum_i I_{1_{\tilde{A}}}^h(i) \leq \frac{2}{\sqrt{2\pi nq(1-q)}} \frac{2n}{\sqrt{2\pi nq(1-q)}} Ent(q) = \frac{4Ent(q)}{2\pi q(1-q)} \leq \frac{8\log(1/q)}{\pi}.$$

Since for $\tilde{A}$, the right hand side of Inequality (13) is $\Theta(1)$, this implies that Proposition 10 is tight, up to a $(\log(1/q))$ multiplicative factor.

4.5   The case $q > 1/2$

In order to show the tightness for $q > 1/2$, we use the *dual function*.

Definition 7 *For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, the dual function $\overline{f} : \{0,1\}^n \to \{0,1\}$ is defined by:*

$$\overline{f}(x_1, x_2, \ldots, x_n) = 1 - f(1 - x_1, 1 - x_2, \ldots, 1 - x_n).$$

It is easy to see that for all $0 < q < 1$,

$$\mathbb{E}_q f = 1 - \mathbb{E}_{1-q} \overline{f}$$

where $\mathbb{E}_q(\cdot)$ denotes expectation w.r.t. the measure μ_q. Similarly, it is easy to show that for all $1 \leq i \leq n$,

$$I_f^q(i) = I_{\overline{f}}^{1-q}(i),$$

where $I_f^q(i)$ denotes influence w.r.t. the measure μ_q. Therefore, the dual functions of the functions described above show the tightness of our results for $q > 1/2$.

References

[1] H. Bauer, Measure and Integration Theory, Walter de Gruyter, New York, 2001.

[2] W. Beckner, Inequalities in Fourier Analysis, *Annals of Math.* **102** (1975), pp. 159–182.

[3] M. Ben-Or and N. Linial, Collective Coin Flipping, in *Randomness and Computation* (S. Micali, ed.), Academic Press, New York, 1990, pp. 91–115.

[4] A. Bonamie, Etude des Coefficients Fourier des Fonctiones de $L^p(G)$, *Ann. Inst. Fourier* **20** (1970), pp. 335–402.

[5] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, and N. Linial, The Influence of Variables in Product Spaces, *Israel J. Math.* **77** (1992), pp. 55–64.

[6] P. Frankl, The Shifting Technique in Extremal Set Theory, in *Surveys in Combinatorics* (C.W. Whitehead ed.), Cambridge University Press, Cambridge, 1987, pp. 81–110.

[7] E. Friedgut, Boolean Functions with Low Average Sensitivity Depend on Few Coordinates, *Combinatorica*, **18(1)** (1998), pp. 27–35.

[8] E. Friedgut, Influences in Product Spaces: KKL and BKKKL Revisited, *Combin. Probab. Comput.*, **13(1)** (2004), pp. 17–29.

[9] E. Friedgut and G. Kalai, Every Monotone Graph Property Has a Sharp Threshold, *Proc. Amer. Math. Soc.* **124** (1996), pp. 2993–3002.

[10] T.E. Harris, A Lower Bound for the Critical Probability in a Certain Percolation Process, *Proc. Cambridge Phil. Soc.* **56** (1960), pp. 13–20.

[11] H. Hatami, Decision Trees and Influence of Variables over Product Probability Spaces, *Combinatorics, Probability and Computing* **18** (2009), pp. 357–369.

[12] J. Kahn, G. Kalai, and N. Linial, The Influence of Variables on Boolean Functions, Proc. 29-th Ann. Symp. on Foundations of Comp. Sci., pp. 68–80, Computer Society Press, 1988.

[13] N. Keller, Lower Bound on the Correlation Between Monotone Families in the Average Case, *Advances in Applied Math.*, to appear. Available on-line at: http://www.ma.huji.ac.il/nkeller.

[14] D.J. Kleitman, Families of Non-Disjoint Subsets, *J. Combin. Theory* **1** (1966), pp. 153–155.

[15] G. A. Margulis, Probabilistic Characteristics of Graphs with Large Connectivity, *Problems Info. Transmission* **10** (1977), pp. 174–179.

[16] E. Mossel, R. O'Donnell and K. Oleszkiewicz, Noise Stability of Functions With Low Influences: Invariance and Optimality, *Annals of Math.*, to appear.

[17] K. Oleszkiewicz, On the Non-Symmetric Khinchine-Kahane Inequality, in *Stochastic Inequalities and Applications* (E. Giné, C. Houdré, and D. Nualart, Eds.), Birkhäuser, 2003, pp. 157–168.

[18] M. Talagrand, Isoperimetry, Logarithmic Sobolev Inequalities on the Discrete Cube, and Margulis' Graph Connectivity Theorem, *Geometric and Functional Analysis* **3** (1993), pp. 295–314.

[19] M. Talagrand, On Russo's Approximate Zero-One Law, *Ann. of Probab.* **22** (1994), pp. 1576–1587.

[20] M. Talagrand, How Much are Increasing Sets Positively Correlated?, *Combinatorica* **16** (1996), no. 2, pp. 243–258.

[21] M. Talagrand, On Boundaries and Influences, *Combinatorica* **17** (1997), pp. 275–285.